DOCUMENT RESUME

ED 191 903 TM 800 527

AUTHOR Fitzgerald, Thomas P.; Clark, Richard M.
TITLE Reactive and Nonreactive Assessment of a Staff Development Program.
PUB DATE May 78
NOTE 23p.; Paper presented at the Annual Meeting of the New England Educational Research Association, (Sturbridge, MA, May 1-3, 1978).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Administrator Attitudes; *Body Language; *Inservice Teacher Education; *Interaction Process Analysis; *Program Evaluation; Reading Consultants; *Semantic Differential; *Staff Development; Teacher Attitudes

ABSTRACT

The reactions of participants to the presentation sessions of a staff development training program were assessed using participant questionnaires and non-verbal metabehavior estimates of participant attitudes. Fifty reading specialists and administrators participated in a two-session workshop dealing with reading material and methods. Each session was evaluated, using a five-point Likert scale, by a participant questionnaire containing three general areas: application of session information to district programs; reaction to the tone established by presenter; and reaction to the organization of session. Additionally, two trained observers assessed the metabehavior of a random sample of nine participants during each session. The metabehavior rating scale reflected the following: body language including head, eyes, mouth, hands, posture and legs; and a seven-point semantic differential including attitude, evaluation, and personality. Reliability of the rating scales was confirmed by an independent evaluation of photographs of selected participants taken during each session. In general, the data indicate: (1) an overall positive response to the sessions; (2) the greatest score differences appeared with eye focus and hand position; and (3) the total mean scores represent positive observations of the metabehavior of the target group. A means for validating metabehavior measures is discussed. (RL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED191903

Reactive and Nonreactive Assessment of A Staff Development Program

Thomas P. Fitzgerald
New York State Education Department

Richard M. Clark
State University of New York at Albany

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
T. Fitzgerald

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Annual Meeting of the New England Educational Research Organization, Sturbridge, Massachusetts, May 1-3, 1978.

TM 800527

Introduction

Staff development programs in schools have continued despite financial, organizational and methodological problems. With less time and financial resources available, the arrangement and duration of such sessions are affected. Now the need for more effective and efficient training programs has increased. Two factors influence the perceived quality of training sessions for teachers: relevancy of the session to the needs of the participants and the appropriateness of the mode of presentation. Methodologically, assessment of these and other factors related to staff development programs have frequently remained superficial.

Staff programs tend not to be rigorously evaluated (Moburg, 1972). Fitzgerald and Marino (1978) stressed the need for a feedback system of evaluation to assist program organizers in refining objectives, content and presentation modes. Frequently, the sole evaluation instrument employed required participants to relate their reactions to the presenter and to the content of the session after the program terminated. More attention should be given to pre-assessment of needs (Ingersoll, 1975), content for session (Otto & Erickson, 1973) and evaluation strategies (Cohen, 1976). As the available resources within educational units are reduced, pressure mounts for more responsible and productive staff development programs.

One major obstacle to improvement rested in the evaluation scheme that was employed. Heavy dependence on verbal reactions of participants may not have generated the accurate feedback information needed to improve programs. Sechrest (1969) emphasized that such verbal

statements are an important manifestation of attitudes, but may not give investigators a fully accurate assessment. Cambell and Stanley (1963) cautioned that the use of questionnaires may precipitate an interaction between testing procedures and the experimental conditions. In other words, the design and phrasing of a questionnaire may outline the expected and acceptable responses rather than encouraging accurate impressions. Centra (1973) reported that teacher ratings on questionnaires are generally positively skewed.

Purpose

The purpose of this study was to investigate the use of observational instruments in evaluating staff training programs. More accurate information may result if questionnaires are combined with non-verbal, metabehavior estimates of participant attitudes (Wilson, 1977). With this in mind, the study focused on the use of body language and semantic differentials as second evaluation tools. More precisely, the study investigated the relationship between the reactive measure of participant questionnaires and non-verbal measures of metabehaviors; namely, body languages and semantic differentials. Secondly, this study addressed the issue of whether observation of metabehavior is similar if completed during the training session as compared with those completed after the session from photographs.

Evaluating Staff Training

Staff training programs have most frequently incorporated attitude questionnaires and other reactive measures as their major

evaluation tools. However, reactive measures may pre-structure attitudes (Wilson, 1977) or interact with the experimental variables (Cambell and Stanley, 1963), thereby reducing the quality and accuracy of the evaluation scheme. As an example, questionnaires may reflect personality and social bias. A halo effect may result from participants unwilling to negatively evaluate sessions and presenters. Conversely, a negative effect may be produced from participants unable to make positive statements because of personal bias. The question arises whether verbal reaction by themselves best represented true perceptions of participants (Sommer, 1967).

Some writers have advanced the principles and techniques employed in ethnographic research as having a potential value in educational research (Strang, 1962; Walcott, 1975; Wolf & Tymitz, 1976). Ethnographers have profitably used the case study, nonreactive, unobtrusive, maturalistic approaches for years in social science research. Walcott (1975) described these approaches as committed to viewing and interpreting social context as seen through the interplay of people and events. Sechrest (1969) suggested that naturalistic measures should be used in conjunction with other tools. Application of unobtrusive and nonreactive measures in educational research are now beginning to be explored. However, efforts to include such approaches in school experimentations required readjustments. For example, the use of participant observers and informants, some frequently used ethnographic tools, are difficult to apply since roles are clearly defined in schools (Walcott, 1975).

The observation and interpretation of non-verbal behavior has been extensively employed in social science research. Semantic differentials and non-verbal behavior have been used to interpret attitudes (Church & Insko, 1965; Pelto, 1970) as has posture and gesture analyses (Sechrest, 1969). Smith-Hanen (1977) studied arm and leg movements while Argyle (1972) described the use of facial expression, looking and head tilt in the analysis of body language.

Duncan (1969) posed the question: if human movements are expressive, how closely do observers agree on their meaning? He also stated that body language is a re-presentation; behavior is then a second form of the inner mode of feelings.

The cues or stimuli employed in this study as focal behaviors are the six most accepted body movement areas (Argyle, 1972; Sechrest, 1969; Schlosberg, 1952). The use of the head, mouth, eyes, hands and posture were readily accessable for observation. The use of leg positions for interpretation was discarded since the workshop participants were seated behind tables.

The semantic differential was constructed with nine semantic opposites with a seven-point scale between each pair. Each of the nine pairs represented one of three categories: attitudes, personality or evaluation.

In-service Training

Staff training has taken many forms. Naumann-Etienne (1977) categorized in-service training into four forms: research and development; social interaction; problem-solving; and advacacy. The interaction of verbal exchanges and attitudes exist not only between student

and teacher (Feldman, 1976), but between workshop presenter and audience. The workshop presentations used in this study might best be categorized as advacacy since both presenters demonstrated potentially useful reading methodology.

The use of observers has been criticized in supervision of teachers (Samph, 1976). However, the observers in this study attempted to operate as participant observers (Becker & Geer, 1957) and, as such, were involved in the workshop activities as members of the audience.

Each session in which observations were completed were one-and-a half hours in duration and included lectures, deomonstration and verbal interaction.

Within this social and instructional context, participants evaluated each session using a traditional reactive instrument and were observed for metabehavior as representative of their attitudes.

Method

Fifty reading specialists and administrators participated in workshop dealing with reading material and methods. The workshop consisted of two sessions by different presenters and different content. Each session was evaluated by a participant questionnaire containing three general areas: application of session information to district programs; reaction to the tone established by presenter, and reaction to the organization of session. Each of these three areas were represented by four positive statements to which the participants agreed or disagreed using a five-point likert scale (one = strongly agree to, five = strongly disagree)(Masters, 1974). Each attendee completed the questionnaire for each session.

To assess participant's metabehavior, a sample of nine participants were randomly selected for each session. Two trained observers analyzed each of the nine member target group during two separate time blocks at five-minute intervals. Therefore, each member was critically observed twice, once by each observers within the one-hour time frame of the session.

The metabehavior rating scale reflected two main elements: body language and a semantic differential. The observation scheme for body language included six interpretable body parts: head (tilting, facing), eyes, mouth, hands, posture and legs (Schlosbery, 1952; Sommer, 1967). A seven-point semantic differential was constructed under three headings: attitude, evaluation and personality (Figure 1). The data generated during the sessions from the observers constituted two evaluations of each sub-group member (N-9) for body language and semantic differentials. A different nine-member sample was identified for the second session.

During each session, four photographs were taken, one every fifteen minutes. The audience was told that publicity material was needed for the newsletter which participants received on a regular basis. These photographs were shown to three pairs of observers trained in the use of the two metabehavior scales. Each pair evaluated three members of the target group to determine the reliability of the rating scales.

Analysis of Data

The past session evaluation by the full audience (N=41) and the target group (N=9) generated fifty response sets for each session

in three areas of application, tone, and reaction. Table 1 depicts a mean-score comparisons between groups and sessions. Since the scale ranged from strongly agree (1) to strongly disagree (5), the lower scores indicate a more positive attitude by participants. The results indicate that session two was more highly regarded by both groups in each category. The target group produced comparable scores to the full audience. Each area was consistantly evaluated by both groups from the most favorable received (tone) to the least well received (reaction). The differences between the full audience and the target group ranged from .06 to .2, with .09 total difference for session one and .11 for session two. In general, the data indicates an overall positive response to the sessions and small differences between the full audience's and the target group's response patterns.

During the sessions, the metabehavior of the target group (N=9 per session) was analyzed through the use of a body language scale and a semantic differential scale. The use of such scales have been advocated in different research projects (Hein, 1976; O'Sullivan & Guilford, 1975; Sindell, 1969). Body language, as reflected in the five elements of head, eyes, mouth, hands and posture, may be viewed as learned behavior which may be decoded by observers in accordance with the rules learned from experience (Duncan, 1969). Mean ratings (Table 2) of eighteen observations by two observers of the target group members reveals a more positive reaction to the second session. Comparing each of the eight scores for the sessions, seven higher scores were recorded during session two, only head tilt differed. In this scale, the high scores (1-5) represented a positive reaction. The greatest score differences appeared with eye focus and hand position which may indicate that

these body parts most clearly demonstrate attitudes or are most easily recorded.

The use of a semantic differential scale has been employed for attitude surveys (Nolen & Archamboult, 1977). Such scales have been demonstrated reliable (LaVoir & Bentler, 1974) and useful for staff evaluation (McDowell, 1974). The data displayed in Table 3 represents the meaning rating on a seven-point scale from negative (1) to positive (7). The pattern of positive to negative opposites and negative to positive opposites were alternated for the observers, but reinterpreted for the table so that high scores indicated positive observations. The data revealed that session two again produced more positive metabehaviors, but the differences were slight. The nine polar opposites may be grouped into attitudes, evaluations and personality reactions. Again, a consistent pattern of observations emerged from the most positive (evaluation) to the least positive behaviors (attitudes) for both sessions. Overall, the total mean scores represent positive observations of the metabehavior of the target group.

One of the main questions addressed in this study is whether participants' reactions correlated with their metabehavior as depicted by body language and semantic differential. High positive correlations would indicate that similar manifestations are being evaluated. The correlations (Table 4) between participants' evaluations and metabehavior ranged from a low of .20 to a high of .86. The highest total correlations were produced during the second session, possibly emphasizing that during more productive presentations, the metabehavior is accentuated. In both sessions, lower correlations were realized when

the semantic differential was used. Such data might be expected since the semantic pairs represent less obvious and less concrete characteristics. The correlation between body language and semantic differential was lower (.50) for the first than the second session (.87).

The observations reported to this point represented judgments made by observers during the sessions. To investigate the possible application of metabehavior analysis following training sessions, photographs were taken at the session and later shown to a group (N=6) of trained observers. Pairs of these observers focused their analysis on three members of the target group. With three pairs, all nine members were again analyzed in terms of their metabehavior. A Kendall Rank-Order Correlation (Edwards, 1967) was computed (Table 5) from the rankings. The data revealed that these two scales of metabehavior generate reliable data and such data may be obtained from photographs of participants. Post-session observations have an advantage in that the observers are not bias by the content of the session, but are at a disadvantage since they are unable to evaluate the participants' moving body language.

As a result of limiting these observers to focus photographs of members of the target group, the metabehavior ratings average approximately one-point lower estimates compared to those produced during session one. These lower scores were true consistently for both body language and semantic differential.

Discussion

In this study, emphasis has been placed on two means of assessing the reactions of participants to staff development training

programs. The means of assessment employed were significantly correlated, but the two sets of measures do not appear to be redundent. A key question concerns the appropriate way to validate the measures employed in this study.

Validation of these measures is only possible if one considers the broader purposes and assumptions of a staff development training program. Such a program can be divided into four parts and evaluation efforts can focus within or between any of these parts. These parts are:

(1) Preparation - during which those responsible for the staff development activity assess the needs of those who will attend the activity, develop the materials and methods to be employed, state the objectives of the activity, arrange the physical setting, etc. Evaluation activity at this level may involve judgments of the needs assessment procedures, the quality of the objectives selected, or the efficiency of the administrative procedures employed.

(2) Presentation - during which material or ideas are presented to those engaged in the staff development activity. This material is assumed to be directed toward some change in the educational practice of the participants. Evaluation at this stage is often directed to participants' evaluation of the merits of the ideas presented, the practicality of these ideas, and the adequacy of the

information presented in assisting one to move toward the idea. In this study, body language was proposed as a means of assessing participants' reactions.

(3) Implementation - during which it is assumed that successful preparation and presentation will lead to the use of the new idea in the educational setting. Evidence of such use may include written records, teacher testimony, or observed teacher behavior.

(4) Effectiveness - during which efforts are made to see when ideas are implemented, different learning or behavioral outcomes are found among those who were expected to be affected. At this stage, evaluation frequently involves analysis of pupil test scores or measures of changed attitudes.

The two sets of measures presented in this study are both taken during the presentation stage of staff development activity. One means of validating these measures is to establish relationships of participant behavior or reaction at the level of presentation to behavior at the level of implementation. One would then ascertain whether which of the measures employed in this study was most predictive of the degree of implementation, or whether a combination of measures is better than a single means of assessment. At a later stage of evaluation, the two measures might be related to effectiveness measures.

Conclusion

In this study, a participant in a staff development program responded to semantic differential items related to attitudes, personality of presenter, and evaluation of the session. In addition, body language assessments were made of nine of the participants. Inter-rater reliability measures of body movement for three pairs of raters ranged from .68 to .96. Body language measures and semantic differential measures were significantly related. A means by which these two alternative measures might be validated is discussed.

TABLE 1

Mean Rating[(a)] of Audience (N=41) Compared to Target Groups (N=9) for Two Presentations

Evaluation Areas	Presentations 1 Aud.	1 Target	2 Aud.	2 Target
Application	2.6	2.4	1.8	1.61
Tone	2.3	2.05	1.7	1.61
Reaction	3.7	3.88	2.08	2.02
Total	2.87	2.78	1.86	1.75

(a) Scores range from positive attitudes (1) to negative attitudes (5)

TABLE 2

Mean Ratings[a] of Body Language Observations of Target Groups (N=9) for Two Presentations

	Presentations	
Body Language	1	2
head tilt	4.25	3.83
facing	4.25	4.66
eye focus	3.75	4.52
mouth	3.50	3.83
hand position	2.75	3.50
activity	3.25	3.66
posture position	4.25	4.66
attentiveness		
Total	3.78	4.15

[a]Scores range from negative (1) to positive (5)

TABLE 3

Mean Ratings from Observations Using Semantic Differential[a] Scale for Two Presentations

Metabehaviors	Presentations 1	2
Evaluation	5.58	5.66
acceptance	5.50	6.32
friendliness	5.75	5.16
understands	5.50	5.50
Attitudes	4.48	4.48
involvement	4.23	4.50
purposefulness	4.25	4.77
responsiveness	5.00	4.16
Personality	4.50	4.83
expressiveness	5.25	6.50
attitude	4.00	3.66
attentiveness	4.25	4.34
Totals	4.85	4.99

[a] Scores ranged from strongly positive (7) to strongly negative (1)

TABLE 4

Correlations[a] Among Metabehavior Observations and Participant's Evaluations

Metabehavior	Participant Evaluations			
	Application	Tone	Reaction	Total
Session one				
Body Language	.44	.70	.64	.82
Semantic Differential	.68	.57	.32	.52
Session two				
Body Language	.51	.84	.91	.86
Semantic Differential	.20	.47	.66	.55

[a]Pearson-Product Moment Correlations

TABLE 5

Mean Rankings and Correlations[a] for Body Language[b] and Semantic Differentials[c] from Pairs of Observers (N=6)

	Observer Pairs					
Metabehavior	1	2	3	4	5	6
Body Language	3.74	3.70	3.92	3.83	3.33	3.75
Correlations	.96		.84		.68	
Semantic Diff.	3.20	3.01	2.88	3.44	2.88	3.72
Correlations	.83		.70		.56	

[a] Kendall's Rank - Order Correlations

[b] Body Language ranks from negative (1) to positive (7)

[c] Semantic Differential ratings ranged from strong negative (1) to strong positive (7)

FIGURE 1

Observation Scale for Metabehavior

Body Language	negative to positive
1. head (related to speaker)	
a) tilting	1 2 3 4 5
b) facing	1 2 3 4 5
2. eyes	
a) toward speaker (idle to purposeful)	1 2 3 4 5
b) away from speaker (idle to purposeful)	1 2 3 4 5
3. mouth (expression)	1 2 3 4 5
4. hands	
a) position (gesture) (idle to purposeful)	1 2 3 4 5
b) activity (idle to purposeful)	1 2 3 4 5
5. posture	
a) attentiveness	1 2 3 4 5
b) position	1 2 3 4 5

Semantic Differential								negative to positive
rejection	1	2	3	4	5	6	7	acceptance
hostile	1	2	3	4	5	6	7	friendly
confused	1	2	3	4	5	6	7	understands
passive	1	2	3	4	5	6	7	involved
placid (face)	1	2	3	4	5	6	7	expressive (face)
distracted	1	2	3	4	5	6	7	engrossed
inferiority (attitude)	1	2	3	4	5	6	7	superiority (attitude)
idle	1	2	3	4	5	6	7	purposeful
unresponsive (manner)	1	2	3	4	5	6	7	responsive (manner)

References

Argyle, Michael. Non-verbal communication in human social interaction. In Non-Verbal Communication (Hinde, ed.). Cambridge University Press, London, 1972.

Becker, H. and B. Geer. Participant observation and interviewing: A comparison. Human Organization. 16, 3, 1957, 28-32.

Cambell, D. and J. Stanley. Experimental and Quasiexperimental designs for research on teaching. In Handbook on the Research in Teaching (Gage, ed) Rand McNally, Chicago, 1963, 171-246.

Carrado, T. and T. Fitzgerald. Ethnography and evaluation: Not-so-strange bedfellows. Paper presented at the Annual Convention of the Northeastern Educational Research Association, Ellenville, New York, October 1977.

Centra, J. Self-ratings of college teachers: A comparison with student ratings. Journal of Educational Measurement. 10, 4, 1973, 287-295.

Church, J. and C. Insko. Ethnic and six differences in sexual values. Psychologia, 8, 1965, 153-157.

Cohen, Karen. Some workable evaluation strategies. Today's Education. 1976, 60-95.

Duncan, S. Non-verbal communication. Psychological Bulletin. 72, 2, 1969, 118-137.

Edwards, A. Statistical Methods. Holt, Rinehart and Winston, Inc. New York, 1967.

Feldman, R. Nonverbal disclosure of teacher deception and interpersonal affect. Journal of Educational Psychology. 68, 6, 1976, 807-816.

Fitzgerald, T. and R. Clark. Process evaluation for in-service training. Reading Improvement.

Fitzgerald, T. and J. Marino. Evaluating in-service staff development in the teaching of reading. Educational Technology. 18, 1978, 49-51.

Hein, G. The science of watching and wondering. The Urban Review, 9, 4, 1976, 242-262.

Ingersoll, G. Assessing in-service training needs through teacher responses. Paper presented at the meeting of the American Educational Research Association, Washington, D. C., April 1975.

References

LaVoir, A. and P. Bentler. A short-form measure of the expanded seven-dimensional semantic space. Journal of Educational Measurement 11, 1, 1974, 65-66.

McDowell, E. The semantic differential as a method of teacher evaluation. Journal of Educational Research. 68, 9, 1975, 330-332.

Masters, J. The relationship between number of responses categories and reliability of likert-type questionnaires. Journal of Educational Measurement. 11, 1, 1974, 49-53.

Moburg, L. In-service teacher training in reading. Newark, Delaware: Reading Information Series. International Reading Association, 1972.

Naumann-Etienne, M. A systematic approach to in-service education. Texas Tech. Journal of Education. 4, 1, 1977, 19-28.

Nolen, W. and F. Archambault. Some issues in semantic differential construction. Paper presented at the meeting of the New England Research Organization, Manchester, N.H., May 1977.

O'Sullivan, M. and J. Guilford. Six factors of behavioral cognition: Understanding other people. Journal of Educational Measurement. 12, 4, 1975, 255-271.

Otto, W. and L. Erickson. In-service education to improve reading instruction. Reading Aids Series. International Reading Association, Newark, Delaware, 1973.

Pelto, P. Anthropological Research, The Structure of Inquiry. Harper and Row. New York, 1970.

Samph, T. Observer Effects on Teacher Verbal Classroom Behavior. Journal of Educational Psychology. 68, [illegible], 1976, 736-741.

Schlosberg, H. The description of facial expressions in terms of two dimensions. Journal of Experimental Psychology. 44, 1952, 229-237.

Sechrest, L. Nonreactive assessment of attitudes. In Naturalistic Viewpoints in Psychological Research (Williams and Raush, eds.) Holt, Rhinehart and Wilson, New York, 1969.

Sindell, P. Anthropological approaches to the study of education. Review of Educational Research. 39, 5, 1969, 593-605.

Smith-Hanen, S. Effects of non-verbal behaviors on judged levels of counselor warmth and empathy. Journal of Counseling Psychology, 24, 2, 1977, 87-91.

References

Sommer, R. Small group ecology. Psychological Bulletin, 67, 1967, 145-152.

Strang, R. Reactions to research on reading. The Education Forum. 26, 1962, 187-192.

Wilson, S. The use of ethnographic techniques in educational research. Review of Educational Research. 47, 1977, 245-265.

Wolcott, H. Criteria for an ethnographic approach to research in schools. Human Organization. 34, 1975, 111-127.

Wolf, P. and B. Tymitz. Ethnography and reading: matching inquiry mode to process. Reading Research Quarterly. 12, 1976, 5-11.